International Journal of Electrical and Computer Engineering (IJECE)
Vol. 14, No. 5, October 2024, pp. 5984~5997
ISSN: 2088-8708, DOI: 10.11591/ijece.v14i5.pp5984-5997

Exploring the frontiers of trajectory outlier detection: an in-depth review and comparative analysis

Sana Chakri[1], Naoual Mouhni[2], Faouzia Ennaama[1]
[1]LAMIGEP, Moroccan School of Engineering Sciences, Marrakech, Morocco
[2]GL-ISI Team, Department of Informatics, Faculty of Science and Techniques, Moulay Ismail University, Meknes, Morocco

Article Info

Article history:

Received May 1, 2024
Revised Jul 10, 2024
Accepted Jul 17, 2024

Keywords:

Autoencoder
Cluster analysis
Deep learning
Outliers
Pattern extraction

ABSTRACT

This paper provides a review and comparative analysis of trajectory outlier detection methods. It presents the definition of outliers in trajectory data and the existing types to further examine the advanced approaches. Basic steps for detecting an outlier, which include data preprocessing, feature extraction, modeling, and similar, have been presented. Moreover, advanced methods such as autoencoders and the use of deep learning for outlier detection have been explored. In the end, this paper evaluates the techniques and compares them using common metrics, mainly focusing on the techniques based on autoencoders or deep learning. It covers applications in real life and practice along with any limitations, challenges, and perspective ideas for the future. Ultimately, it can be a useful resource for expanding the understanding of domain researchers and practitioners.

Corresponding Author:

Sana Chakri
LAMIGEP, Moroccan School of Engineering Sciences
Marrakech 40000, Morocco
Email: chakri.Sana@gmail.com

1. INTRODUCTION

Detecting outliers in trajectories is a fundamental problem in many fields such as traffic monitoring, fleet management, and intrusion detection. The analysis of a trajectory as a sequence of points that describes the movement of an object in space and time helps identify abnormal behavior or rare and significant events. Detecting outliers in trajectories is critical to ensuring safety, optimizing resources, preventing accidents, and promoting informed decision-making. Trajectories represent the movement of objects in space and time, and identifying outliers on trajectories aims to detect abnormal behaviors or rare and significant events. Trajectory anomalies can be defined as typical behaviors or movements compared to normal conditions. These may include sudden changes in direction, unusual speeds, or deviations from expected spatial patterns. Figure 1 illustrates the case of expected trajectory outliers that can be extracted between two regions. Outliers are expected to be sub-trajectories that have some neighbors nearby, while normal trajectories have more neighbors nearby [1].

The initial part of this extensive research focuses on the significant issue of identifying outliers in trajectories, which is crucial in various fields such as traffic monitoring, fleet management, and intrusion detection [2]. By examining trajectories as sequences of points that represent an object's movement in both space and time [3], [4], it becomes feasible to detect abnormal behavior or rare, significant events. The ability to detect outliers within trajectories is essential for ensuring safety, optimizing resources, preventing accidents, and facilitating informed decision-making. Our main emphasis in this study is on the detection of outliers within trajectories, with a specific focus on recent advancements that encompass classical methods, machine learning algorithms, and deep learning-based techniques. We explore various definitions of trajectory outliers, traditional detection methods, associated challenges, constraints, and innovative approaches rooted in machine learning principles [5]. Additionally, we investigate the recent progress in utilizing deep learning for identifying trajectory outliers and engage in insightful discussions regarding applications and potential areas for further research.

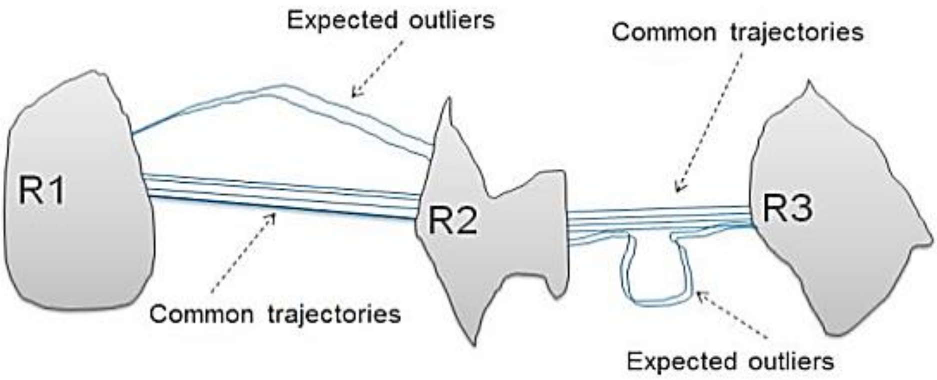

Figure 1. Outliers between two regions

The primary goal of this paper is to provide a comprehensive overview of trajectory outlier detection methods, highlighting the advantages, disadvantages, and performance metrics of different approaches. We seek to clarify current limitations and challenges in trajectory outlier detection using existing methods, identify opportunities for future research endeavors, and offer practical insights through case studies. We aim to present a detailed overview of trajectory outlier detection, covering classical methods, recent advancements in machine learning techniques, and the incorporation of deep learning methodologies, thereby contributing to the advancement of anomaly detection protocols in trajectory data analysis [6].

The remaining sections of this paper are structured as follows: section 2 will discuss the common steps to detect outliers in trajectories, while section 3 will cover advanced techniques for outlier detection in trajectories. In section 4, an evaluation and comparison of outlier detection techniques will be provided, and section 5 will focus on applications of outlier detection in trajectories. Lastly, section 6 will address limitations and challenges, and the final section will present the conclusion.

2. COMMON STEPS TO DETECT OUTLIERS

2.1. Trajectory data pre-processing

Before starting the identification of outliers, it is typically essential to preprocess the trajectory data to prepare it for analysis. This preprocessing stage encompasses various steps, including data normalization to eliminate any discrepancies in scale between trajectory variables, as well as trajectory sampling to address issues related to variable density [7]. Trajectory data pre-processing passes through a set of micro-points which are:

- Data cleaning: this step involves handling missing values, removing outliers, and resolving inconsistencies within the trajectory data. Techniques such as imputation and statistical methods are commonly employed to accomplish this task.
- Coordinate system conversion: convert trajectory data into a standardized coordinate system for consistency and ease of analysis.
- Noise reduction and smoothing: to enhance the quality of the trajectory data, noise reduction techniques are applied. These techniques, such as moving average or Gaussian filtering, help to smoothen the trajectory and eliminate any irregularities present.
- Data normalization: the data is normalized using the min-max normalization technique, which brings the trajectory values within the range of 0 to 1. This normalization process is crucial as it eliminates any scale differences that may exist between the variables [8].
- Trajectory sampling: to select a representative subset of trajectories or trajectory points, while maintaining the integrity of the dataset, trajectory sampling techniques are employed. Importance measures or clustering methods are commonly used to achieve this. For example, Mirge *et al.* [9] proposed a sampling method based on the use of importance measures to select the most significant trajectory points.

2.2. Extraction of trajectory features

Once the data has been pre-processed, it is necessary to extract trajectory features that will serve as a basis for outlier detection. This can include measures based on distance, density, or specific indicators related to trajectory behavior. Noulas *et al*. [10] identified various features from trajectory data to identify outliers. One of the methods used involved extracting points, lines, and significant regions [11]. The process commences with defining the scope and determining specific trajectory features based on the research objectives or problem requirements. Subsequently, it computes distance-based features like total distance traveled, distance between consecutive points, or distance from reference points using appropriate distance metrics. Additionally, density-based features are extracted, including point density, cluster density, or area coverage using density metrics like kernel density estimation or spatial grid-based counting techniques. Duration-based features are also considered by analyzing timestamps associated with trajectory points to calculate features related to total duration, segment durations, or time spent in specific regions. Furthermore, speed and acceleration indicators are captured by extracting features such as average speed, maximum speed, changes in speed over time, average acceleration, maximum acceleration, or sudden changes in acceleration. Trajectory direction and turning behavior are quantified through features like average direction, direction deviation, or the number of turns, utilizing techniques like angular deviation or angle calculations. Moreover, contextual features relevant to the application or research context, such as road types, traffic conditions, or weather conditions, are extracted. Finally, feature extraction techniques, including simple calculations, statistical measures, or specialized algorithms like Fréchet distance or dynamic time warping, are applied to extract the trajectory features.

2.3. Modeling approaches for outlier detection

Once the features have been extracted, various modeling approaches can be employed to identify outliers in trajectories [12]. One commonly used method is mixture models, which are statistical models that represent normal trajectories and identify outliers by comparing trajectory points with the models. These models assume that the trajectory data consists of a mixture of multiple distributions, with outlier trajectories represented by a distinct distribution. Unsupervised mixture models like Gaussian mixture models (GMM) or Dirichlet process mixture models can be employed for this purpose. For instance, in [12], unsupervised mixture models were applied to identify anomalies in vehicle trajectory data. Another method is statistical techniques, often used for outlier detection in trajectory data. These techniques involve analyzing the statistical properties of trajectory features to detect deviations from the normal pattern. This may include metrics such as z-scores: This paper aims to solve two enduring challenges in existing trajectory similarity measures: computational inefficiency and the absence of the ‘uniqueness’ property that should be guaranteed in a distance function: $dist(X, Y) = 0$ if and only if $X = Y$, where X and Y are two trajectories. In this work, we present a novel approach utilizing a distributional kernel for trajectory representation and similarity measurement, based on the kernel mean embedding framework. It is the very first time a distributional kernel is used for trajectory representation and similarity measurement. Our method does not rely on point-to-point distances which are used in most existing distances for trajectories. Unlike prevalent learning and deep learning approaches, our method requires no learning. We show the generality of this new approach in anomalous trajectory and sub-trajectory detection. We identify that the distributional kernel has i) a data-dependent property and the ‘uniqueness’ property which are the key factors that lead to its superior task-specific performance, and ii) runtime orders of magnitude faster than existing distance measures, Mahalanobis distance, or statistical tests like the Grubbs' test or the Dixon's test. The third approach discussed is ensemble approaches, which combine multiple outlier detection models to improve the accuracy and reliability of the detection process. These approaches may entail aggregating the outcomes of various models, each utilizing a different algorithm or feature set, to make a final decision on outlier detection. Boosting, bagging, or random forest methods are commonly employed in ensemble approaches. The final approach mentioned is the trajectory-specific model, where certain outlier detection techniques are specifically tailored for trajectory data analysis. These models take into account the temporal and spatial characteristics of trajectories and integrate domain-specific knowledge. Examples include moving object trajectory outliers (MOTO) and trajectory outlier detection (TOD) methods. It is crucial to consider the specific characteristics of the data when selecting the appropriate modeling approach [13].

3. ADVANCED TECHNIQUES FOR OUTLIER DETECTION

In this segment, we shall delve into various sophisticated methodologies employed for identifying outliers in trajectories. As previously stated, outliers, also known as anomalous values, refer to trajectory points that exhibit substantial deviation from the overall pattern of the data. Precise identification of these outliers holds utmost importance in guaranteeing the accuracy of analyses and models reliant on trajectories.

We shall thoroughly investigate these advanced techniques, encompassing their mechanisms, benefits, and constraints, to present a comprehensive outline of the most efficient approaches for outlier detection in trajectories. A summary of all the approaches discussed will be provided in Table 1.

Table 1. Advanced techniques for outlier detection in trajectories

Approach	Use Case	Advantages	Drawbacks	Algorithms used
Nearest neighbor outlier detection (NNOD)	– Data streams with high velocity and continuous flow of data – Real-world dataset including text documents and images	– Simplicity and robustness – Efficiency: significantly faster than alternative methods – Flexibility: applicable to any arbitrary distance measure – Applicable to various domain	– Dependence on parameters such as the number of neighbors and the deviation threshold – Fine-tuning of parameters may be required for optimal performance – Computational complexity	– k-nearest neighbors' algorithm (k-NN) – Local Outlier Factors (LOF) – Exemplar-based Nearest Neighbor Outlier Detection (ENNOD)
Trajectory shape-based outlier detection (TSOD)	– Identify outliers in trajectory data by focusing on the shape of the trajectories – Behavior analysis in sports	– Flexibility: Used with different types of trajectory data, such as vehicles, people, and animals – Consideration of spatial and Geometric context	– Need to select appropriate similarity measures or modeling functions. – Significant data reduction is required to produce reasonable results – Dependence on feature extraction	– Dynamic Time Warping (DTW) – Fourier Analysis-Based Approaches
Spatio-temporal outlier detection (STOD)	– Used to identify aberrant values in trajectory data, considering both spatial and temporal dimensions – Urban traffic trajectory data	– Advanced approach used to identify anomalies in trajectory data, considering both spatial and temporal dimensions	– Complexity of trajectory data, spatiotemporal dynamics, and the presence of noise	– Clustering methods – Prediction models – Density-based approaches – DBSCAN
Autoencoders for outlier detection (AEOD)	Identify outliers in the case of: – Surveillance video data – Urban traffic trajectory data	– Reduces the calculation amount, making it suitable for real-time detection scenarios. – Achieves high accuracy, with precision and recall rates exceeding 95%.	– Requires a pre-training step – Performance heavily depends on the quality of the training data – Sensitivity to certain types of anomalies or specific data distributions	– Deep spatial-temporal autoencoders (DSTAE) – Variational Autoencoders (VAE)
Deep learning for outlier detection (DLOD)	– Advanced technique that has gained increasing popularity in outlier detection in trajectory data	– Ability to extract important features from data	– Requires large amounts of data for training – Challenges related to data availability and result interpretability	– Convolutional neural networks (CNN) – Recurrent neural networks (RNN)

3.1. Nearest neighbor outlier detection

The nearest neighbor outlier detection (NNOD) technique is a popular approach in the field of trajectory analysis. This method is based on the principle that outliers in trajectories can be identified by examining their distance from their nearest neighbors in the feature space. One commonly used method in nearest neighbor-based outlier detection is the k-nearest neighbors (k-NN) algorithm [12], [14]. In this algorithm, the k nearest neighbors of a given trajectory point is identified and used to measure the point's proximity and deviation from its surroundings. The main focus of the approach used is on the development and analysis of an efficient algorithm for computing minimax k-NN and its application to outlier detection. The minimax distance measure is a distance measure that seeks the minimum largest gap among all different routes between two objects. The minimax distance between objects i and j is computed as:

$$D\_M\_M\_i,j = min\_\{r \in R\_ij\ (O)\}\,\{max\_\{1 \leq l < |r|\}\ D\_r(l)r(l+1)\}$$

where $R_{ij}(O)$ is the set of all routes between i and j, and each route r is specified by a sequence of object indices. The minimax distance is obtained by finding the minimum of the maximum edge weights along all the different routes between the two objects. To compute the minimax k-NN of a new object v, an incremental approach is used. The algorithm iteratively extends the set of partial neighbors of v by adding the next minimax nearest neighbor at each step, based on the minimum distance to the set of already selected neighbors. Overall, the minimax distance measure captures the underlying data geometry by focusing on the largest gap among different routes, and the algorithm efficiently computes the minimax k-NN for a given test object. Various versions of this algorithm have been developed to improve the detection of outliers in trajectories. One such variant is the local outlier factors (LOF) algorithm. The LOF algorithm assesses the abnormality of a trajectory point by comparing its local density with that of its neighbors. By calculating the degree of outlyingness for each data point, the LOF algorithm can identify local outliers within the dataset.

Alghushairy *et al*. [15] proposed a hybrid approach combining LOF and the k-NN algorithm for outlier detection in motion trajectories. The key components of the LOF algorithm include the k-distance, which measures the distance between a data point and its $k^{\text{-th}}$ nearest neighbor, and the k-NN themselves, which form the set of k data points closest to a given data point. Additionally, the algorithm incorporates the reachability distance (RD) to measure the local density between two data points, and the local reachability distance (LRD), which represents the average ratio of the local reachability density of a data point and its k-NN. Each data point is assigned an LOF score by the LOF algorithm, which is then used to determine if it is an outlier. This algorithm is effective in identifying local density and detecting local outliers within the dataset. However, it is important to note that the LOF algorithm may have a long execution time and can be sensitive to the minimum points value. In summary, the LOF algorithm operates by evaluating the local density of data points and identifying outliers based on their deviation from the local density of their neighbors.

Another advanced technique used for outlier detection is the exemplar-based nearest neighbor outlier detection (ENNOD) approach. This method involves selecting trajectory points that significantly differ from their nearest neighbors in terms of features. Sengupta and Das [16] proposed an approach based on ENNOD to detect outliers in trajectory data. This method is rooted in the scientific principles of proximity analysis and anomaly detection. It involves systematically identifying the nearest neighbors for each data point within a dataset, using distance metrics like Euclidean distance or other similarity measures. By considering the local context of each data point and evaluating its relationship with neighboring points, the method calculates a measure of outlier-ness based on the distances to its nearest neighbors. Data points that exhibit significant deviations in distance compared to their neighbors are flagged as potential outliers. The classification of data points as outliers is further determined by applying a threshold based on their distances to their nearest neighbors. This approach finds applications in various domains such as anomaly detection, fraud detection, and quality control, where the identification of rare and unusual patterns is crucial. Overall, exemplar-based nearest neighbor outlier detection effectively utilizes the concept of nearest neighbors and local context to identify data points that deviate significantly from their local neighborhood [17], [18], indicating potential outliers. The use of this technique offers advantages such as simplicity and robustness. However, it is important to note that this method may also have limitations [19], particularly in terms of its dependence on parameters like the number of neighbors and the deviation threshold. Fine-tuning of these parameters may be necessary to optimize the performance of outlier detection in trajectory data.

3.2. Trajectory shape-based outlier detection

In contrast to conventional outlier detection techniques that depend on distance metrics, trajectory shape-based outlier detection (TSOD) concentrates on examining the shapes of trajectories. Instead of only taking into account the spatial or temporal distance among trajectory points, TSOD evaluates the overall structure and configuration of the trajectory itself [20]. The TSOD method typically includes the extraction of shape-based attributes from the trajectories. These attributes may involve properties like curvature, speed, acceleration, and changes in direction along the trajectory. By capturing these shape-based attributes, TSOD aims to pinpoint unique patterns or irregularities that could signify outliers. TSOD might employ shape-based distance metrics to compare the extracted attributes among trajectories [21]. These metrics gauge the similarity or dissimilarity of trajectory shapes, enabling the detection of outliers based on deviations from standard trajectory patterns. The method frequently entails statistical analysis to assess the shape-based attributes and distance metrics. This analysis could encompass techniques such as clustering, density estimation, or hypothesis testing to pinpoint trajectories that significantly differ from the norm. Through the use of shape-based attributes and distance metrics, TSOD identifies trajectories that display unusual shape characteristics or deviate notably from the anticipated patterns. These trajectories are marked as outliers, indicating potential anomalies or abnormal behavior within the dataset.

One prevalent method in shape-based outlier detection is the dynamic time warping (DTW) algorithm [22]. DTW is a method that gauges the similarity between two trajectories by aligning them temporally and computing the optimal alignment distance. In [9], dynamic time warping (DTW) is utilized for outlier detection in trajectory data. The primary aim of employing the DTW algorithm in this research is to investigate the variability in car-following behavior among drivers and the diverse situation-dependent behavior of drivers within a single trip. This algorithm is implemented to estimate the time-varying parameters of a car-following model using detailed vehicle trajectory data. It is employed to determine the optimal alignment between two sets of time-series data [23], which serves as an approximation of the stimulus-response relationship for a driver following another vehicle. This approximation is then leveraged to deduce the time-varying parameters of the car-following model for a driver over the observation period. Furthermore, Fourier analysis-based methods have been utilized for detecting outliers in trajectories based on shape. For instance, the utilization of Fourier series to represent trajectories enables the extraction of frequency features and the identification of outliers through spectral deviations. Noulas *et al.* [10] introduced a Fourier analysis-based technique for outlier detection in trajectories. The primary goal of this proposed method is to organize trajectories to discern and categorize both typical and atypical behaviors exhibited by objects like aircraft and ships. The strategy involves utilizing feature vectors to concisely represent the significant details in trajectories, encompassing fundamental information like total distance covered and the distance between starting and ending points, along with geometric characteristics associated with the convex hull properties, trajectory curvature, and overall distance geometry. The objective is to leverage these features to recognize trajectories resembling a prototype, organize a repository of numerous trajectories, and pinpoint anomalies. It is important to highlight that employing the trajectory shape-based outlier detection method poses specific challenges. A primary obstacle is the necessity to opt for suitable similarity metrics or modeling functions that align with the particular data and trajectory types under consideration. Therefore, meticulous adjustment and validation of these metrics and functions are imperative to achieve precise outcomes in shape-based outlier detection within trajectories. To sum up, the TSOD method represents an advanced strategy for detecting irregular values in trajectory data. Techniques based on DTW and Fourier analysis have proven effective in shape-based outlier detection within trajectories. Nonetheless, it is crucial to carefully select the appropriate parameters and similarity metrics to secure dependable and meaningful outcomes in shape-based outlier detection within trajectories.

3.3. Spatio-temporal outlier detection

The spatiotemporal outlier detection (STOD) technique is a sophisticated method used to detect abnormal values in trajectory data, taking into account both the spatial and temporal aspects. The primary objective of this approach is to identify spatiotemporal anomalies that differ from the overall trajectory behavior [11]. STOD leverages the fusion of spatial and temporal information within trajectory data to pinpoint outliers. It incorporates the geographical locations (spatial dimension) and the corresponding time of visit (temporal dimension) within the trajectory data. By considering both dimensions, STOD captures the movement patterns and behaviors of objects or individuals over time and space. The technique involves extracting pertinent features from trajectories, such as spatial coordinates, timestamps, speed, acceleration, direction changes, and other spatiotemporal attributes, which collectively provide a comprehensive representation of the trajectory data [24]. STOD employs distance measures that account for both spatial and temporal dimensions to evaluate the similarity or dissimilarity between trajectories. By comparing the spatio-temporal features of trajectories, STOD can identify deviations from typical movement patterns, indicating potential outliers. Statistical analysis is often employed to assess the spatiotemporal features and distance measures [25]. Utilizing these features and measures, STOD identifies trajectories that exhibit unusual movement patterns or deviate significantly from the expected spatiotemporal behaviors, flagging them as outliers. STOD finds applications in various fields such as transportation, urban planning, environmental monitoring, and surveillance. It aids in recognizing unusual movement behaviors and pinpointing suspicious or irregular activities of individuals within a specific area over a period. The technique of spatio-temporal outlier detection offers a holistic method for pinpointing abnormal values in trajectory data by taking into account both the spatial and temporal aspects of movement patterns. One prevalent method in spatio-temporal outlier detection involves utilizing clustering techniques.

Clustering works towards grouping similar trajectory points and singling out clusters that deviate significantly from the rest of the dataset. Zhang *et al.* [26] introduce an approach that leverages trajectory clustering and peak densities for spatiotemporal outlier detection. Clustering is utilized for grouping similar trajectories, while peak density is employed to pinpoint isolated and rare trajectories. An alternative method involves the utilization of prediction models, which forecast expected trajectories by analyzing past trajectory behavior and pinpointing deviations from these predictions. Wang *et al.* [18] introduced a technique that utilizes a recurrent neural network model to predict trajectories and identify spatiotemporal outliers. A recurrent neural network-based temporal prediction model was utilized to anticipate trajectory movements and compare them with actual observations to identify outliers. In addition, density-based approaches have also been employed for spatio-temporal outlier detection. These approaches leverage the spatial and temporal distribution of trajectory data to identify trajectory points that are either highly isolated or located in densely populated regions. Bai *et al.* [27] proposed a density-based analysis method for detecting spatiotemporal outliers, utilizing the density-based spatial clustering of applications with noise (DBSCAN) algorithm to compute the density of trajectory points and identify low-density points, which are considered outliers. It is worth noting that the detection of spatio-temporal outliers presents challenges due to the intricate nature of trajectory data, spatio-temporal dynamics, and the presence of noise. Therefore, a careful selection of similarity measures, clustering parameters, and prediction models is crucial in order to obtain accurate results in spatio-temporal outlier detection [12].

3.4. Autoencoders for outlier detection

Autoencoders (AEOD) have emerged as a recently advanced technique for detecting outliers in trajectory data. These neural network models are capable of learning and representing input data in a concise and precise manner through an encoding-decoding structure [28]. By training the autoencoder on normal trajectory data, it becomes adept at capturing the underlying patterns and features of the trajectories. Throughout the training phase, the autoencoder acquires the ability to encode the input trajectory data into a lower-dimensional representation, commonly referred to as a latent space or bottleneck layer [29]. This encoding process aims to capture the fundamental characteristics of the input trajectories while simultaneously reducing their dimensionality. The autoencoder then proceeds to decode this representation of lower dimensionality, effectively returning it to its original form within the input space. Once the autoencoder has been trained using regular trajectory data, it becomes capable of reconstructing or generating trajectories. When presented with new or previously unseen trajectory data, the autoencoder endeavors to reconstruct the input trajectories by relying on the patterns it has learned. If the input trajectory closely aligns with the learned patterns, the resulting reconstruction error is expected to be minimal. Conversely, if the input trajectory significantly deviates from the learned patterns, the reconstruction error will be substantial. This reconstruction error serves as a metric for measuring the dissimilarity between the input trajectory and the learned representations. Trajectories that yield high reconstruction errors are regarded as potential outliers or anomalies. This is due to their failure to conform to the learned patterns and characteristics of normal trajectories. In the realm of trajectory data, autoencoders can effectively identify outliers by leveraging their ability to capture intricate patterns and dependencies within trajectories. By acquiring a concise representation of regular trajectory data and detecting deviations through reconstruction errors, autoencoders offer a promising approach for detecting anomalies within trajectory data. This functionality makes them particularly suitable for identifying abnormal movement patterns, unexpected stops, irregular speeds, or unconventional routes within trajectory datasets. Numerous recent studies have demonstrated the efficacy of autoencoders in outlier detection within trajectories. For example, Li *et al.* [30] proposed a method for spatio-temporal outlier detection based on autoencoders. An autoencoder was employed to acquire the representation of normal trajectories, and the identification of outliers was achieved by comparing observed trajectories with those reconstructed by the autoencoder. The research paper outlines the application of an autoencoder within the two-stream deep spatial-temporal auto-encoder (two-stream DSTAE) framework for detecting anomalies in surveillance videos. Specifically, the autoencoder is utilized in both the spatial and temporal streams to extract appearance characteristics and motion patterns, respectively. In the spatial stream, an autoencoder is employed to extract appearance characteristics from the original RGB video frames.

Through unsupervised learning, the autoencoder learns the regular surveillance videos, effectively addressing the issue of imbalanced data between positive and negative samples. The reconstruction error generated by the autoencoder is utilized to differentiate abnormal events, with low reconstruction errors indicating regular videos and high reconstruction errors indicating abnormal videos. Similarly, the temporal stream employs an autoencoder to extract motion patterns from continuous optical flow frames. By learning features from the optical flow frames, the autoencoder in the temporal stream contributes to the extraction of motion characteristics. The results of their study demonstrate that the utilization of autoencoders enables accurate detection of spatio-temporal outliers in trajectory data. Additionally, the research also explores the use of variational autoencoders (VAE) in spatio-temporal outlier detection. A VAE is a variant of autoencoders that models the latent distribution of data and allows for generating new data instances. VAEs are a type of autoencoder that not only reconstructs input data but also models the underlying distribution of the data in a latent space. This means that VAEs can learn the probability distribution of the input data and generate new data instances that are similar to the training data. In the context of spatio-temporal outlier detection, VAEs can capture the complex patterns and correlations present in the data, making them suitable for identifying outliers in spatio-temporal datasets. By learning the distribution of normal data, VAEs can identify data points that deviate significantly from this learned distribution, thereby flagging them as potential outliers. Zhang *et al.* [31] proposed a VAE-based approach for outlier detection in trajectories.

Their method leverages a VAE to learn the distribution of normal trajectories and generate synthetic trajectories. The model calculates the difference between an original trajectory and the trajectory generated by the model using three different distance metrics: vertical distance ($d \perp$), parallel distance ($d \parallel$), and angular distance ($d\theta$).

- Vertical distance ($d\ \perp$): This metric measures the vertical distance between trajectory segments. It captures the positional deviation between the original trajectory and the trajectory generated by the model.
- Parallel distance ($d\ \parallel$): The parallel distance calculates the parallel distance between trajectory segments, capturing the positional deviation in the direction of the trajectory.
- Angular distance ($d\theta$): This measurement assesses the angular variance between segments of a trajectory, reflecting the divergence in trajectory direction. The dissimilarity between the initial trajectory and the produced trajectory is articulated as a blend of three distance metrics, each with designated weight coefficients. Through the examination of various distance and deviation elements, the model effectively encompasses the distinctions between the original trajectory and the model-generated trajectory. Subsequently, the recorded trajectories are juxtaposed with the generated trajectories, pinpointing outliers characterized by significant dissimilarity.

3.5. Deep learning for outlier detection

Deep learning (DLOD) has become increasingly popular in outlier detection in trajectory data due to its advanced capabilities [32]. One of its key strengths lies in its ability to automatically extract hierarchical features from trajectories. This is particularly important given the complex spatial and temporal patterns that are often found in trajectory data. Deep learning models, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), are particularly effective in processing trajectory data and capturing spatial and temporal dependencies. CNNs excel at extracting spatial features [33], while RNNs are adept at capturing temporal dependencies and sequential patterns within trajectories [34]. In the context of outlier detection, deep learning models are trained on large amounts of normal trajectory data. This enables them to effectively identify outliers by evaluating the deviation of new or unseen trajectories from the learned patterns. Moreover, deep learning models can be combined with unsupervised learning techniques to perform anomaly detection in trajectory data, further enhancing their effectiveness in this field. One of the benefits of utilizing deep learning for outlier detection lies in its flexibility to handle various types of trajectory data, such as global positioning system (GPS) trajectories, movement trajectories, and spatiotemporal data from different sources. This adaptability enables deep learning models to be effectively utilized across a broad spectrum of trajectory datasets in diverse domains like transportation, urban planning, environmental monitoring, and surveillance. deep learning, a subset of artificial intelligence, utilizes deep neural networks to autonomously acquire intricate data representations. This ability to extract essential features from data positions it as a promising technique for detecting anomalies in trajectories. The sensitivity of deep learning to specific anomalies or unique data distributions stems from its capacity to automatically learn intricate data representations, allowing it to identify anomalies that deviate from the learned representations. In practical terms, this sensitivity empowers deep learning models to detect anomalies that might not be easily identifiable using conventional methods. By learning directly from the data, deep learning models can adjust to and pinpoint anomalies that are characteristic of particular data distributions or display distinct patterns, thus enhancing their sensitivity to such anomalies in trajectory data. Numerous recent studies have showcased the efficacy of deep learning in outlier detection within trajectories. For instance, Taylor *et al*. [35] introduced a deep learning-based method for spatiotemporal outlier detection in trajectory data.

The application of a convolutional neural network (CNN) was employed to automatically extract spatial and temporal features from trajectories. By comparing observed trajectories with normal trajectories learned by the CNN model, outliers were identified based on dissimilarity. The results demonstrated that deep learning techniques facilitated accurate detection of spatio-temporal outliers in trajectory data. Moreover, the utilization of recurrent neural networks (RNN) has also been investigated for spatio-temporal outlier detection. RNNs are specifically designed neural networks that excel in handling sequence data, making them well-suited for modeling trajectories. Rintoul *et al*. [36] proposed an RNN-based approach for detecting spatiotemporal outliers in maritime trajectory data. Their approach involved employing an RNN-based prediction model to anticipate expected trajectories and identifying points that exhibited significant differences between predictions and actual observations as outliers. To detect outliers in the weighted neighborhood information network (WNIN), the proposed approach utilized a customized Markov random walk method. The Markov random walk process was employed to determine the inlier degree for each object by calculating the stationary distribution vector. This vector quantified the inlier degree of objects in the dataset with mixed-valued attributes. The Markov random walk method was specifically designed to capture long-range correlations within the dataset and calculate the outlier score of each object based on the network representation provided by the WNIN. This method was tailored to handle the unique characteristics of datasets with mixed-valued attributes and served as a crucial component of the proposed outlier detection approach. It is important to note that the use of deep learning in outlier detection typically necessitates a substantial amount of training data, which can pose challenges in cases where trajectory data is limited. Additionally, the selection of the neural network architecture is crucial in achieving optimal results.

3.6. Commonly used evaluation metrics

Evaluation metrics are of utmost importance when it comes to comparing and evaluating outlier detection techniques in trajectory data. These metrics enable us to assess the performance of methods and quantify their accuracy, recall, specificity, and other significant measures. Numerous commonly used evaluation metrics have been proposed in the literature to cater to the specific requirements of outlier detection. Among these metrics, one that stands out as widely used in outlier detection is the area under the curve (AUC) [37]. In the context of outlier detection, AUC serves as a measure of how effectively a model can differentiate between normal and outlier data points. The calculation of AUC involves plotting the true positive rate (sensitivity) against the false positive rate (1-specificity) for various threshold values. The resulting AUC value ranges from 0 to 1, with a higher value indicating superior performance. An AUC of 0.5 signifies a model that performs no better than random chance, while an AUC of 1 represents a perfect model. In the context of outlier detection, a higher AUC suggests that the model excels at distinguishing between normal and outlier data points. This implies that the model is more proficient at identifying outliers and minimizing false positives (normal data points incorrectly classified as outliers) and false negatives (outliers incorrectly classified as normal).

AUC was employed by Zhang *et al.* [26] to evaluate the performance of various outlier detection methods in drone trajectory data. The findings of their study revealed that approaches utilizing deep learning and unsupervised learning techniques achieved higher AUCs, thereby demonstrating their efficacy in detecting anomalies. Another commonly employed metric, Precision, plays a crucial role in evaluating the accuracy of a model's predictions, particularly in the context of outlier detection. It quantifies the proportion of correctly identified outliers among all the points labeled as outliers by the model. In the realm of outlier detection, minimizing false positives is often of utmost importance. False positives occur when normal data points are erroneously classified as outliers. This can result in unnecessary alerts or actions, which can be both costly and disruptive. Precision proves particularly valuable in such scenarios as it focuses on the model's ability to accurately identify true outliers while minimizing false positives. For instance, if a model designates 100 points as outliers, with 90 of them being true outliers and 10 being false positives, the precision would be 90%. This signifies that 90% of the points identified as outliers by the model are indeed true outliers.

In a study by Zhang *et al* [38], on outlier detection in vehicle trajectory data, precision served as the primary metric for evaluating method performance. The results demonstrated that certain density-based methods achieved high levels of precision, making them viable options in scenarios where minimizing false positives is critical. In addition to AUC and precision, outlier detection also utilizes other evaluation metrics like recall, F1-score, accuracy, and F-measure [39]. Table 2 showcases the standard metrics employed in the field of outlier detection. These metrics evaluate different aspects of method performance, including the ability to detect genuine outliers and reduce classification errors.

Table 2. Common metrics for outlier detection evaluation

Metric	Use Case
AUC	Evaluates the ability of a model to distinguish between normal and outlier data points.
Precision	Measures the proportion of true outliers among all points identified as outliers by the model, useful for minimizing false positives.
Recall	Measures the proportion of true outliers that are correctly identified by the model, useful for capturing all actual outliers.
F1 Score	Combines precision and recall to provide a single metric that balances both false positives and false negatives.
Accuracy	Measures the overall correctness of the model's predictions, but may not be suitable if the data is imbalanced.
Specificity	Measures the proportion of true negatives that are correctly identified by the model, useful for minimizing false alarms.
Mean squared error	Measures the average of the squares of the errors between actual and predicted values, useful for regression-based outlier detection.
Mean absolute error	Measures the average of the absolute differences between actual and predicted values, useful for regression-based outlier detection.

3.7. Outlier detection application fields

Outlier detection in trajectories is utilized in a wide range of fields. For instance, within the realm of traffic management and road safety, the identification of abnormal vehicle behaviors can be facilitated through outlier detection. An illustration of this is seen in the work of Chen *et al.* [40], who introduced a technique for detecting anomalies in vehicle trajectories by focusing on abnormal driving patterns, allowing for the identification of dangerous situations or traffic violations.

In the realm of maritime and aerial surveillance, the identification of outlier detection in trajectories holds significant importance in uncovering suspicious or unlawful activities. A notable example is the work of Hu *et al.* [41], who devised a method utilizing generative models to effectively detect abnormal ship trajectories, the identification of suspicious maritime movements that may indicate illicit activities or territorial violations can be facilitated. Additionally, outlier detection in trajectories is employed in healthcare and medical monitoring. For instance, the detection of anomalies in the trajectories of patients with chronic diseases can assist physicians in recognizing alterations in behavior or severe clinical situations, thereby enabling timely intervention and proactive care management. According to Wu *et al.* [42], outlier detection in the trajectories of patients with Alzheimer's disease can be utilized to pinpoint abnormal behaviors. Moreover, within the realm of environmental and wildlife surveillance, the detection of outliers in trajectories plays a crucial role in pinpointing unusual behaviors exhibited by animals or endangered species. As an illustration, Yang *et al*. [43] introduced a novel approach utilizing deep autoencoders to identify outliers in elephant trajectories, facilitating the detection of unusual behaviors like herd separation or territorial loss. Additionally, outlier detection in trajectories shows potential in various fields like logistics, cybersecurity, and financial fraud detection, expanding its applicability and introducing new possibilities for its utilization across different sectors.

3.8. Limitations and challenges

Despite the considerable progress made in outlier detection in trajectories, there are still significant limitations and challenges that must be overcome to achieve precise and dependable anomaly detection. One such challenge is the susceptibility of trajectory data to measurement errors, noise, and data gaps, which can hinder the accurate identification of outliers. Moreover, the voluminous and intricate nature of trajectory data introduces additional complexities, as it frequently includes errors, gaps, and uncertainties, thereby further complicating the process of detecting anomalies [44]. The subjective nature of defining outliers introduces an additional level of intricacy and subjectivity to the task of identification. The definition of outliers may vary depending on the particular field of application. It is essential to grasp the context and needs of the domain in order to accurately define and identify outliers within trajectories [45].

The detection of outliers is further complicated by the dynamic characteristics of trajectories and the temporal progression of movements. Trajectories can display seasonal fluctuations, trends, or intricate patterns as time progresses, resulting in transient outliers that may not necessarily indicate genuine anomalies. It is crucial to consider the temporal dynamics to prevent the identification of false positives [46]. The magnitude and intricacy of trajectory data pose further obstacles. Vast trajectory data containing multiple dimensions and attributes heighten the complexity of their analysis and outlier detection. Moreover, the computationally intensive techniques and substantial computing resources needed for outlier detection in spatiotemporal trajectories contribute to the complexity [47]. Furthermore, the complexity of outlier detection is exacerbated by the high dimensionality of trajectory data and the variability of trajectories.

Detecting abnormal patterns becomes increasingly difficult when dealing with trajectories that contain a large number of data points and variables. Moreover, the dynamic nature of trajectories, which can be influenced by external factors, introduces an additional level of intricacy to outlier detection. The existence of missing or noisy values in trajectory data presents yet another major obstacle. Inaccurate or absent trajectory points resulting from sensor malfunctions or data inaccuracies have the potential to skew the outcomes of outlier detection, thereby compromising the precision and dependability of anomaly detection within trajectory paths [48]. Lastly, the diversity of outliers and trajectories makes it challenging to select an appropriate anomaly detection method. Different techniques may be more effective for specific types of trajectories or anomalies. This necessitates the tailoring of detection methods to match the specific characteristics of the trajectory data and enable accurate detection of anomalies.

4. COMPARATIVE STUDY

The performance comparison among various outlier detection techniques holds great significance in the field of trajectory data analysis. Several research studies have been carried out to assess and contrast the effectiveness of different anomaly detection methods, utilizing metrics such as AUC, precision, recall, and other evaluation criteria. In this comparative analysis, four different algorithms were utilized to examine a dataset containing stock market information. These algorithms represented a wide array of methodologies, including density-based, statistical-based, and deep learning-based approaches. The LOF algorithm, a density-based technique, utilized local density to identify outliers within the dataset. Principal component analysis (PCA), a statistical-based method, was employed to reduce data dimensionality and effectively detect anomalies. Moreover, the study incorporated the use of AutoEncoder (AE), a deep learning-based method, to capture intricate patterns and interactions within the dataset. Additionally, the DeepSVDD algorithm, another deep learning-based approach, was used to model the low-dimensional space where normal data points are located. The dataset, which consisted of stock market data, represented a complex and dynamic environment influenced by various factors. The application of these algorithms aimed to identify anomalies that could indicate potential manipulation, significant news events, or other significant shifts within the stock market. Figure 2 illustrates the outliers detected by each method. To assess the study's effectiveness, key metrics such as accuracy, precision, recall, and F1 score were employed. These metrics offered a comprehensive evaluation of the algorithms' performance in detecting anomalies within the stock market data, facilitating a thorough comparative analysis of their efficacy. This study aimed to provide valuable insights into the application of diverse outlier detection techniques.

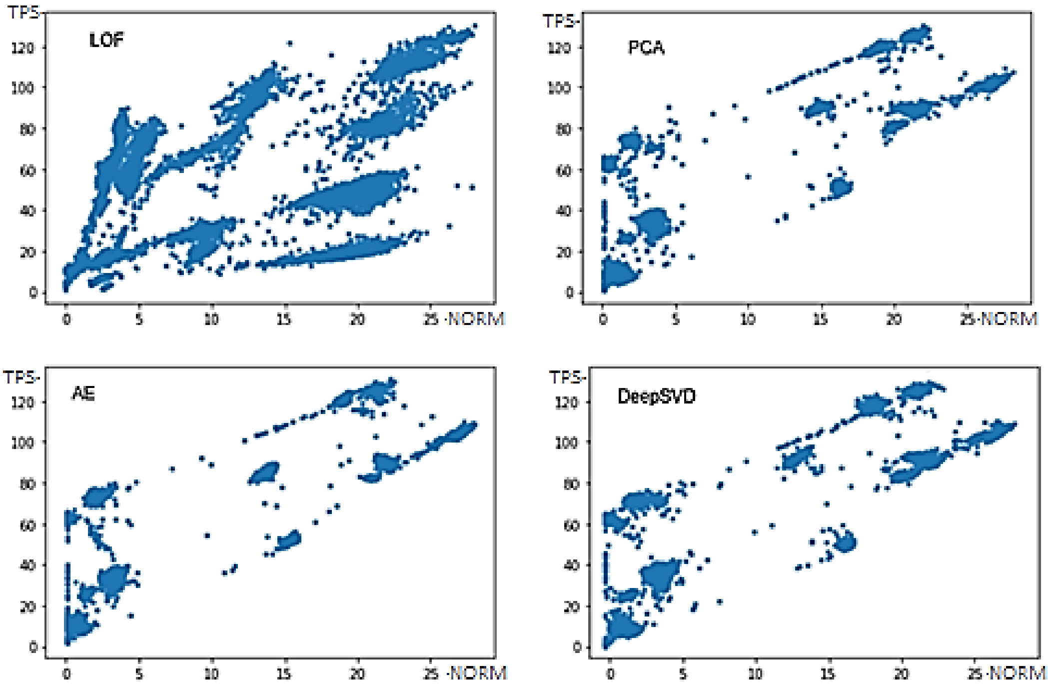

Figure 2. Outliers detected

4.1. Results and discussions

The results displayed in Table 3 provide a detailed overview of the performance metrics for each machine learning model, including accuracy, precision, recall, and F1 score. The recall value for all models indicates that they were able to identify the majority of true positive cases, while the F1 score provides a balanced measure of precision and recall. The results offer valuable insights into the comparative effectiveness of the machine learning models for anomaly detection in the context of stock market data analysis. Based on the findings of the comparative study, several noteworthy observations can be derived. The local outlier factor (LOF) technique exhibits the highest level of accuracy, reaching an impressive 96.42% among the evaluated methods. It also achieves a perfect recall score of 1.0, indicating its effectiveness in identifying all relevant instances within the dataset. Although its precision score of 0.657 suggests the possibility of some false positives, overall, it performs well with an F1 score of 0.764.

Table 3. Metrics for each machine learning model

Method	Accuracy	Precision	Recall	F1 score
LOF	0.9642	0.657	1.0	0.764
PCA	0.9731	0.556	0.98	0.708
AE	0.9856	0.465	0.98	0.628
DeepSVDD	0.9848	0.438	0.98	0.638

Following closely is the PCA method, which attains an accuracy of 97.31%. Despite having a relatively lower precision score of 0.556, it demonstrates a high recall rate of 0.98, resulting in an F1 score of

0.708. The AE method showcases the highest accuracy among the tested techniques, standing at an impressive 98.56%. However, it does exhibit a lower precision score of 0.465 and an F1 score of 0.628, indicating the potential for improvement in terms of precision. The deep support vector data description (DeepSVDD) method achieves an accuracy of 98.48% with a precision score of 0.438 and an F1 score of 0.638. Despite its lower precision score, it maintains a high recall rate of 0.98, suggesting its effectiveness in capturing relevant instances within the dataset.

Through an extensive comparative analysis of outlier detection techniques in trajectory data analysis, valuable insights have been obtained regarding the effectiveness of different machine learning algorithms for detecting anomalies in stock market data. Evaluating methods like LOF, PCA, AE, and DeepSVDD provided a thorough assessment of their performance metrics, highlighting their strengths and limitations in this domain. While each algorithm showed unique capabilities in detecting anomalies, distinct patterns emerged from the findings. LOF demonstrated high accuracy and recall rates, proving its effectiveness in identifying outliers. On the other hand, PCA excelled in reducing data dimensionality but had room for improvement in precision. AE showed high accuracy but needed better precision, indicating a balance between accuracy and precision in its use. Despite a lower precision score, DeepSVDD exhibited strong recall capabilities, emphasizing its ability to capture relevant instances in the dataset.

Comparing these results with previous research showcases the progress in outlier detection methodologies and emphasizes the importance of selecting algorithms tailored to specific data characteristics and research goals. The study's strength lies in evaluating multiple algorithms in a real-world financial dataset, providing detailed insights into their performance metrics and implications for anomaly detection [38]. Furthermore, the differences in precision among the algorithms raise important questions about enhancing detection accuracy while minimizing false positives, suggesting avenues for further research and algorithm enhancement [49]. As trajectory data analysis remains crucial in detecting anomalies in dynamic systems, this study contributes to advancing outlier detection techniques in this field [50].

5. CONCLUSION

To conclude, this investigation into identifying outliers in trajectories highlights the critical significance of continuously advancing algorithmic methodologies to effectively tackle the evolving complexities found in modern data environments. By thoroughly examining various techniques for outlier detection, this study has provided insights into the strengths and limitations of existing approaches, emphasizing the necessity for a diverse range of algorithms to identify anomalies in trajectory data accurately. The future direction of this field relies on the ongoing refinement and development of robust techniques capable of handling the inherent high dimensionality of trajectory datasets. Embracing innovations such as deep learning-based approaches and the integration of diverse data sources, including contextual factors like weather and socioeconomic data, presents a promising frontier for enhancing precision in outlier detection. Additionally, the utilization of semi-supervised learning techniques offers a significant opportunity to improve accuracy in scenarios where labeled anomaly data is scarce, ultimately paving the way for more effective anomaly detection. These advancements not only enhance the effectiveness of outlier detection but also have profound implications across various domains, including security surveillance and transportation management. By uncovering these possibilities for future research and practical implementation, this study acts as a catalyst for propelling the field of outlier detection in trajectories toward increased efficiency and applicability in real-world contexts. As researchers and practitioners continue to explore these avenues, the potential for transformative impacts on anomaly detection methodologies and their practical implications remains vast and promising.

REFERENCES

[1] S. Chakri, S. Raghay, and S. El Hadaj, “Enriching trajectories with semantic data for a deeper analysis of patterns extracted,” in *Advances in Intelligent Systems and Computing*, vol. 552, Springer International Publishing, 2017, pp. 209–218.

[2] J. Gudmundsson, P. Laube, and T. Wolle, “Movement patterns in spatio-temporal data,” in *Encyclopedia of GIS*, Springer US, 2008, pp. 726–732.

[3] J. H. Sullivan, M. Warkentin, and L. Wallace, “So many ways for assessing outliers: What really works and does it matter?,” *Journal of Business Research*, vol. 132, pp. 530–543, Aug. 2021, doi: 10.1016/j.jbusres.2021.03.066.

[4] W. Oueslati, S. Tahri, H. Limam, and J. Akaichi, “A systematic review on moving objects’ trajectory data and trajectory data warehouse modeling,” *Computer Science Review*, vol. 47, Feb. 2023, doi: 10.1016/j.cosrev.2022.100516.

[5] M. Gupta, J. Gao, C. Aggarwal, and J. Han, “Outlier detection for temporal data,” *Synthesis Lectures on Data Mining and Knowledge Discovery*, vol. 5, no. 1, pp. 1–129, Mar. 2014, doi: 10.2200/s00573ed1v01y201403dmk008.

[6] C.-T. Lu, D. Chen, and Y. Kou, “Algorithms for spatial outlier detection,” in *Third IEEE International Conference on Data Mining*, 2003, pp. 597–600, doi: 10.1109/ICDM.2003.1250986.

[7] S. Chakri, S. Raghay, and S. El Hadaj, “Semantic trajectory knowledge discovery: a promising way to extract meaningful patterns from spatiotemporal data,” *International Journal of Software Engineering and Knowledge Engineering*, vol. 27, no. 3, pp. 399–421, Apr. 2017, doi: 10.1142/S0218194017500140.

[8] R. M. Assunção, M. C. Neves, G. Câmara, and C. Da Costa Freitas, "Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees," *International Journal of Geographical Information Science*, vol. 20, no. 7, pp. 797–811, Aug. 2006, doi: 10.1080/13658810600665111.
[9] V. Mirge, S. Gupta, and K. Verma, "A novel approach for mining trajectory patterns of moving vehicles," *International Journal of Computer Applications*, vol. 104, no. 4, pp. 4–8, Oct. 2014, doi: 10.5120/18188-9097.
[10] A. Noulas, S. Scellato, N. Lathia, and C. Mascolo, "Mining user mobility features for next place prediction in location-based services," in *2012 IEEE 12th International Conference on Data Mining*, Dec. 2012, pp. 1038–1043, doi: 10.1109/ICDM.2012.113.
[11] S. Chakri, S. Raghay, and S. El, "Modeling, mining, and analyzing semantic trajectories: the process to extract meaningful behaviors of moving objects," *International Journal of Computer Applications*, vol. 124, no. 8, pp. 15–21, Aug. 2015, doi: 10.5120/ijca2015905542.
[12] M. H. Chehreghani, "K-Nearest neighbor search and outlier detection via minimax distances," in *Proceedings of the 2016 SIAM International Conference on Data Mining*, Jun. 2016, pp. 405–413, doi: 10.1137/1.9781611974348.46.
[13] X. Wu, Y. Liu, X. Zhao, and J. Chen, "STKST-I: an efficient semantic trajectory search by temporal and semantic keywords," *Expert Systems with Applications*, vol. 225, Sep. 2023, doi: 10.1016/j.eswa.2023.120064.
[14] J. Yang, X. Tan, and S. Rahardja, "Outlier detection: How to Select k for k-nearest-neighbors-based outlier detectors," *Pattern Recognition Letters*, vol. 174, pp. 112–117, Oct. 2023, doi: 10.1016/j.patrec.2023.08.020.
[15] O. Alghushairy, R. Alsini, T. Soule, and X. Ma, "A review of local outlier factor algorithms for outlier detection in big data streams," *Big Data and Cognitive Computing*, vol. 5, no. 1, pp. 1–24, Dec. 2021, doi: 10.3390/bdcc5010001.
[16] S. Sengupta and S. Das, "Selective nearest neighbors clustering," *Pattern Recognition Letters*, vol. 155, pp. 178–185, Mar. 2022, doi: 10.1016/j.patrec.2021.10.005.
[17] Y. Zhou, H. Xia, D. Yu, J. Cheng, and J. Li, "Outlier detection method based on high-density iteration," *Information Sciences*, vol. 662, Mar. 2024, doi: 10.1016/j.ins.2024.120286.
[18] Y. Wang, Y. Ma, H. Huang, B. Wang, and D. P. Acharjya, "A split–merge clustering algorithm based on the k-nearest neighbor graph," *Information Systems*, vol. 111, Jan. 2023, doi: 10.1016/j.is.2022.102124.
[19] P. Indyk and H. Xu, "Worst-case performance of popular approximate nearest neighbor search implementations: Guarantees and limitations," *Advances in Neural Information Processing Systems*, vol. 36, pp. 66239–66256, 2023.
[20] Y. Wang, Z. Wang, K. M. Ting, and Y. Shang, "A principled distributional approach to trajectory similarity measurement and its application to anomaly detection," *Daimon*, vol. 79, pp. 865–893, Mar. 2024, doi: 10.1613/jair.1.15849.
[21] D. Xia *et al.*, "An ASM-CF model for anomalous trajectory detection with mobile trajectory big data," *Physica A: Statistical Mechanics and its Applications*, vol. 621, Jul. 2023, doi: 10.1016/j.physa.2023.128770.
[22] A. Giannoula *et al.*, "Exploring long-term breast cancer survivors' care trajectories using dynamic time warping-based unsupervised clustering," *Journal of the American Medical Informatics Association*, vol. 31, no. 4, pp. 820–831, Jan. 2024, doi: 10.1093/jamia/ocad251.
[23] S. Chakri and N. Mouhni, "Time-series neural network predictions using LSTM for clustering and forecasting GPS data of bird immigration," *2024 International Conference on Global Aeronautical Engineering and Satellite Technology (GAST)*, Marrakesh, Morocco Apr. 2024, doi: 10.1109/GAST60528.2024.10520764.
[24] E. O. Eldawy, M. Abdalla, A. Hendawi, and H. M. O. Mokhtar, "Spatio-temporal outlier detection: methods, applications, and research directions," in *Emerging Trends, Techniques, and Applications in Geospatial Data Science*, IGI Global, 2023, pp. 63–79.
[25] D. Hu, L. Chen, H. Fang, Z. Fang, T. Li, and Y. Gao, "Spatio-temporal trajectory similarity measures: a comprehensive survey and quantitative study," *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 5, pp. 2191–2212, May 2024, doi: 10.1109/TKDE.2023.3323535.
[26] P. Zhang, M. Deng, Y. Shi, and L. Zhao, "Detecting hotspots of urban residents' behaviours based on spatio-temporal clustering techniques," *GeoJournal*, vol. 82, no. 5, pp. 923–935, May 2017, doi: 10.1007/s10708-016-9720-4.
[27] X. Bai, Z. Xie, X. Xu, and Y. Xiao, "An adaptive threshold fast DBSCAN algorithm with preserved trajectory feature points for vessel trajectory clustering," *Ocean Engineering*, vol. 280, Jul. 2023, doi: 10.1016/j.oceaneng.2023.114930.
[28] M. N. K. Sikder and F. A. Batarseh, "Outlier detection using AI: a survey," in *AI Assurance: Towards Trustworthy, Explainable, Safe, and Ethical AI*, Elsevier, 2022, pp. 231–291.
[29] G. A. Noghre, A. D. Pazho, and H. Tabkhi, "An exploratory study on human-centric video anomaly detection through variational autoencoders and trajectory prediction," in *2024 IEEE/CVF Winter Conference on Applications of Computer Vision Workshops (WACVW)*, Jan. 2024, pp. 995–1004, doi: 10.1109/WACVW60836.2024.00109.
[30] T. Li, X. Chen, F. Zhu, Z. Zhang, and H. Yan, "Two-stream deep spatial-temporal auto-encoder for surveillance video abnormal event detection," *Neurocomputing*, vol. 439, pp. 256–270, Jun. 2021, doi: 10.1016/j.neucom.2021.01.097.
[31] L. Zhang, W. Lu, F. Xue, and Y. Chang, "A trajectory outlier detection method based on variational auto-encoder," *Mathematical Biosciences and Engineering*, vol. 20, no. 8, pp. 15075–15093, 2023, doi: 10.3934/mbe.2023675.
[32] W. Chen *et al.*, "Deep learning for trajectory data management and mining: a survey and beyond," *arXiv:2403.14151*, Mar. 2024.
[33] S. Yaqoob, S. Cafiso, G. Morabito, and G. Pappalardo, "Detection of anomalies in cycling behavior with convolutional neural network and deep learning," *European Transport Research Review*, vol. 15, no. 1, Mar. 2023, doi: 10.1186/s12544-023-00583-4.
[34] Y. Su, D. Yao, T. Tian, and J. Bi, "Transfer learning for region-wide trajectory outlier detection," *IEEE Access*, vol. 11, pp. 97001–97013, 2023, doi: 10.1109/ACCESS.2023.3294689.
[35] J. Taylor, X. Zhou, N. M. Rouphail, and R. J. Porter, "Method for investigating intradriver heterogeneity using vehicle trajectory data: a dynamic time warping approach," *Transportation Research Part B: Methodological*, vol. 73, pp. 59–80, Mar. 2015, doi: 10.1016/j.trb.2014.12.009.
[36] M. D. Rintoul and A. T. Wilson, "Trajectory analysis via a geometric feature space approach," *Statistical Analysis and Data Mining*, vol. 8, no. 5–6, pp. 287–301, Oct. 2015, doi: 10.1002/sam.11287.
[37] D. Chicco and G. Jurman, "The matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification," *BioData Mining*, vol. 16, no. 1, Feb. 2023, doi: 10.1186/s13040-023-00322-4.
[38] X. Zhang, Y. Zheng, Z. Zhao, Y. Liu, M. Blumenstein, and J. Li, "Deep learning detection of anomalous patterns from bus trajectories for traffic insight analysis," *Knowledge-Based Systems*, vol. 217, Apr. 2021, doi: 10.1016/j.knosys.2021.106833.
[39] B. U. M. Babu, K. G. Babu, and B. T. Krishna, "Trajectory interception classification for prediction of collision scope between moving objects," *Scalable Computing*, vol. 25, no. 2, pp. 1167–1190, Feb. 2024, doi: 10.12694/scpe.v25i2.2621.
[40] C. Chen *et al.*, "Abnormal-trajectory detection method based on variable grid partitioning," *ISPRS International Journal of Geo-Information*, vol. 12, no. 2, Jan. 2023, doi: 10.3390/ijgi12020040.

[41] J. Hu *et al.*, "Intelligent anomaly detection of trajectories for IoT empowered maritime transportation systems," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 2, pp. 2382–2391, 2023, doi: 10.1109/TITS.2022.3162491.
[42] H. Wu, X. Tang, Z. Wang, and N. Wang, "Uncovering abnormal behavior patterns from mobility trajectories," *Sensors*, vol. 21, no. 10, May 2021, doi: 10.3390/s21103520.
[43] Y. Yang, Z. Zhang, W. Mao, Y. Li, and C. Lv, "Radar target recognition based on few-shot learning," *Multimedia Systems*, vol. 29, no. 5, pp. 2865–2875, Jul. 2023, doi: 10.1007/s00530-021-00832-3.
[44] H. Zhang, Y. Luo, Q. Yu, L. Sun, X. Li, and Z. Sun, "A framework of abnormal behavior detection and classification based on big trajectory data for mobile networks," *Security and Communication Networks*, vol. 2020, pp. 1–15, Dec. 2020, doi: 10.1155/2020/8858444.
[45] Z. Wang, G. Yuan, H. Pei, Y. Zhang, and X. Liu, "Unsupervised learning trajectory anomaly detection algorithm based on deep representation," *International Journal of Distributed Sensor Networks*, vol. 16, no. 12, Dec. 2020, doi: 10.1177/1550147720971504.
[46] J. Li, X. Pei, X. Wang, D. Yao, Y. Zhang, and Y. Yue, "Transportation mode identification with GPS trajectory data and GIS information," *Tsinghua Science and Technology*, vol. 26, no. 4, pp. 403–416, Aug. 2021, doi: 10.26599/TST.2020.9010014.
[47] W. Cao, S. Wang, P. N. Landin, C. Fager, and T. Eriksson, "Complexity optimized digital predistortion model of RF power amplifiers," *IEEE Transactions on Microwave Theory and Techniques*, vol. 70, no. 3, pp. 1490–1499, Mar. 2022, doi: 10.1109/TMTT.2021.3122956.
[48] P. Han *et al.*, "Calibrations of low-cost air pollution monitoring sensors for CO, NO2, O3, and SO2," *Sensors*, vol. 21, no. 1, Jan. 2021, doi: 10.3390/s21010256.
[49] H. Pan, Y. Li, and D. Zhao, "Recognizing human behaviors from surveillance videos using the SSD algorithm," *Journal of Supercomputing*, vol. 77, no. 7, pp. 6852–6870, Jan. 2021, doi: 10.1007/s11227-020-03578-3.
[50] J. Wang, X. Guo, and X. Yang, "Efficient and safe strategies for intersection management: a review," *Sensors*, vol. 21, no. 9, Apr. 2021, doi: 10.3390/s21093096.

BIOGRAPHIES OF AUTHORS

Sana Chakri holds a Ph.D. in software engineering from FSTG, UCA Marrakech and an engineering degree in software engineering from FST Mohammedia, UH2 Casablanca. She completed her DUT in computer science from FSTG, UCA Marrakech. Their influential scholarly articles on semantic trajectory knowledge discovery have illuminated moving object behaviors and pattern extraction. Dr. Chakri's expertise in data mining bridges theoretical research and practical applications, addressing real-world challenges with innovative solutions. Their commitment to cutting-edge research in outlier detection systems enriches the scientific community and inspires the next generation of researchers. Dr. Chakri's current research focuses on new methods for outlier detection, aiming to deepen understanding in the field and foster the development of robust outlier detection systems. She can be contacted at chakri.sana@gmail.com.

Naoual Mouhni SMIEEE, holds a Ph.D. in computer science and is currently a professor at FSTE University Moulay Ismail. Previously, she served as the director of research and development at the LAMIGEP multidisciplinary research laboratory at EMSI Marrakech. She has been teaching as a computer science professor at the same institution since 2012. Naoual is a passionate advocate for technology and innovation and holds a variety of certifications that showcase her broad range of skills. She has obtained specialized training in artificial intelligence from the NVIDIA Deep Learning Institute, educational advancements from Honoris University, expertise in big data from IBM, and proficiency in database management from ORACLE. Her research focuses on pushing the boundaries of artificial intelligence, renewable energies, and database technologies. Her commitment to advancing these fields is evident through her ongoing research and contributions to each area. She serves as a senior member of the IEEE and is actively involved in several academic and professional communities. She can be contacted at n.mouhni@umi.ac.ma.

Faouzia Ennaama graduated with a master's degree in electrical engineering and a Ph.D. in electrical engineering in telecommunications and industrial techniques, she obtained her doctorate after conducting her research at the LAPSSI laboratory at Cadi Ayyad University in Marrakech, Morocco. Passionate about facial recognition, computer vision, electronics, and programming, she has published papers in indexed journals and participated in several international conferences. She can be contacted at faouzia.ennaama@ced.uca.ma.